

**TecnoLógicas**  
ISSN-p 0123-7799  
ISSN-e 2256-5337  
Vol. 22, edición especial  
Noviembre de 2019  
pp. 119-132

Artículo de revisión /review Article

## Interacciones basadas en gestos: revisión crítica

### Gestural Based Interactions: Critical Review

Laura Cortés-Rico <sup>1</sup> y  
Giovanny Piedrahita-Solórzano <sup>2</sup>

Recibido: 25 de septiembre de 2019  
Aceptado: 19 de noviembre de 2019

---

#### Cómo citar / How to cite

L. Cortés-Rico, G. Piedrahita-Solórzano, "Interacciones basadas en gestos: revisión crítica", *TecnoLógicas*, vol. 22, pp. 119-132, 2019. <https://doi.org/10.22430/22565337.1512>

© Instituto Tecnológico Metropolitano  
Este trabajo está licenciado bajo una  
Licencia Internacional Creative  
Commons Atribución (CC BY-NC-SA)



- <sup>1</sup> MSc. en Ingeniería de Sistemas y Computación, Universidad Militar Nueva Granada, Cajicá-Colombia, [laura.cortes@unimilitar.edu.co](mailto:laura.cortes@unimilitar.edu.co)
- <sup>2</sup> MSc. en Ingeniería Electrónica, Institución Universitaria Politécnico Grancolombiano, Bogotá- Colombia, [gapiedahita@poligran.edu.co](mailto:gapiedahita@poligran.edu.co)

## **Resumen**

Este artículo presenta una revisión crítica de la interacción humano-computador (HCI) basada en gestos. El gesto, como una forma de comunicación no verbal, ha sido de interés para el área de HCI en la búsqueda de alternativas de interacción entre el humano y la máquina, a través del cuerpo como agente que percibe y actúa en el mundo. La revisión se hizo en las bases de datos de mayor importancia en HCI y en algunas fuentes de literatura académica latinoamericana en el área, e incluye un análisis de la evolución de las interacciones basadas en gestos, el trabajo actual y las perspectivas a futuro. El análisis se desarrolla de forma holística y abarca asuntos técnicos y humanos: psicológicos, sociales y culturales, así como su relación. Este proceso analítico se presenta como una descripción cienciométrica de los resultados de las búsquedas, a fin de exponer el gesto como medio de interacción, las técnicas utilizadas para los diferentes pasos en el proceso de reconocimiento de gestos y las aplicaciones y desafíos de las interacciones basadas en gestos. Como conclusión se formula una serie de preguntas que invitan al lector a pensar en potenciales focos de investigación en las interacciones basadas en gestos.

## **Palabras clave**

Gestos, interacción humano-computador, reconocimiento de gestos, interacciones basadas en gestos.

## **Abstract**

This paper presents a critical review of human-computer interactions (HCI) based on gestures. Gestures, as ways of non-verbal communication, have been of interest in HCI because they make possible the interaction with the machine through the body, as an agent that perceives and acts in the world. The review was carried out in the most critical databases in HCI, as well as some Latin-American academic sources, and included an analysis of the evolution of gesture-based interactions, current work, and future perspectives. The article is carried out holistically, considering both technical and human issues: psychological, social, and cultural, as well as their relationships. We present this analytical process as a scientometric description of the search results, the description of the gesture as a means of interaction, the techniques used for the different steps in the gesture recognition process, and the presentation of the applications and challenges of gesture-based interactions. It concludes through a series of questions that invite the reader to think about potential research focus on gesture-based interactions.

## **Keywords**

Gestures, human-computer interaction, gesture recognition, gesture based interactions.

## 1. INTRODUCCIÓN

La palabra gesto tiene su origen en el latín *gestus*, que hace referencia a una forma de comunicación no verbal basada en lenguajes corporales. Los gestos son expresiones o movimientos faciales, de las manos o cualquier parte del cuerpo, a través de los cuales se manifiestan pensamientos, sentimientos o estados de ánimo. Su propósito es el de intercambiar eficientemente un mensaje entre quien emite el gesto y quien lo interpreta.

Adicionalmente, el latín *gestus* tiene relación con *gerere*, que a su vez significa “llevar a cabo”; de allí la relación de la palabra gesto con otras como ‘gestionar’, ‘gestar’ o ‘gerencia’ [1]. En español, gesto también se utiliza para referir a un conjunto de acciones que tienen un significado o intencionalidad particular [2].

En el área de la interacción humano-computador (HCI) se investiga la utilización de gestos como medio de comunicación con dispositivos de cómputo, en las que el cuerpo es un agente principal en la comunicación. Adicionalmente, en algunas ocasiones se utiliza para evaluar la experiencia del usuario al enfrentarse a ciertas interacciones; por ejemplo, para estimar las emociones que genera una interacción de acuerdo con los gestos que hace el usuario [3]. La investigación posee dos enfoques relevantes: uno técnico — ¿cómo construir interfaces que permitan a los dispositivos computacionales reconocer gestos de las personas? [4], [5], ¿cómo representar computacionalmente los gestos? [6]-[8], ¿de qué manera las máquinas pueden representar información a través de gestos por los cuales las personas sientan empatía? [9]-[11]—; y otro humano (psicológico, cultural, social) — ¿cómo los gestos se transforman por la inclusión de dispositivos de cómputo en la cotidianidad? [12], [13], ¿qué factores intervienen en la aceptación/apropiación social de las interacciones basadas en gestos? [13], ¿cuál es la influencia cultural

en los gestos y cómo esa influencia interfiere en la estandarización de los procesos de reconocimiento y representación computacional de gestos? [14]-[16]—.

Este artículo presenta una revisión crítica de literatura académica y de divulgación científica enfocada en las interacciones basadas en gestos, que responde algunas de las preguntas enunciadas anteriormente. La sección 2, metodología, corresponde al proceso llevado a cabo para hacer la revisión, los criterios de búsqueda, la agrupación y la sistematización de la información recolectada. En la sección 3, se discuten los resultados obtenidos en el proceso de búsqueda, los cuales son presentados en subsecciones a manera de relatos y esquemas, con referencias que invitan al lector a profundizar sobre los temas de interés; asimismo, se formulan preguntas que buscan indagar acerca de los potenciales temas de investigación, desafíos, riesgos y aplicaciones de las interacciones basadas en gestos. El artículo finaliza con las conclusiones, que sintetizan los hallazgos y ofrecen una reflexión crítica sobre las interacciones humano-computador basadas en gestos, la cual pone en evidencia el interés de la comunidad científica en este tema y la necesidad de abordarlo desde perspectivas técnicas, sociales y culturales. Así mismo, finaliza con el planteamiento de la posibilidad de estudiar otras direcciones de interacción, como los gestos que una máquina puede hacer a un humano.

## 2. METODOLOGÍA

La revisión presentada en este artículo parte del interés en desarrollar métodos y herramientas de interacción entre las personas y los computadores, en los que la frontera entre lo material —objetos, ambiente, sujetos— y lo inmaterial —bits, datos, representaciones abstractas— sea

difusa. Con frecuencia este tipo de interacciones son referidas en áreas de HCI como interfaces tangibles de usuario [17], [18], interfaces orgánicas (OUI) [19] o interfaces naturales de usuario (NUI) [20].

En todos los casos, es común encontrar al cuerpo como agente principal, que percibe y actúa.

Igualmente, cuando se hace una búsqueda del papel del cuerpo en procesos de comunicación e interacción entre humanos y entre el humano y la máquina, frecuentemente se asocia a los gestos, entendidos como medio no verbal de comunicación. En razón a esto, se decidió hacer una revisión que permitiera condensar algunos de los hechos más relevantes de las interacciones humano-computador mediadas por gestos.

La búsqueda se desarrolló con base en el proceso metodológico presentado en la Fig. 1. La primera etapa consistió en definir una versión inicial de las ecuaciones de búsqueda y establecer cuáles serían las principales fuentes de información. La segunda, en buscar dichas ecuaciones en las fuentes de información establecidas, especialmente desde un enfoque científico/académico —bases de datos—, pero también desde uno de divulgación —magazines, redes sociales—.

La tercera etapa involucró actividades de revisión y análisis de los resultados de la búsqueda. Esta revisión se hizo en términos *cientiométricos* para determinar la cantidad de resultados obtenidos, las principales conferencias y revistas de publicación, los autores con mayor cantidad de textos, los años de publicación y las palabras frecuentes en los resultados, entre otros. Adicionalmente, y en un continuo ir y venir con la siguiente etapa metodológica, se esbozó una aproximación inicial a los textos: lectura de título, resumen y palabras clave. La cuarta etapa metodológica consistió en el establecimiento de los criterios para seleccionar o no un texto, hacer únicamente una lectura rápida (adicionar títulos de secciones, gráficos, tablas y conclusiones) o leerlo con mayor detalle.

Finalmente, se sistematizó información a partir de la generación de relaciones y preguntas que derivaron de las etapas previas. La metodología es iterativa, en el sentido de que las etapas están estrechamente conectadas y la sistematización de la información resultante generó continuamente nuevas ecuaciones de búsqueda y nuevas necesidades en torno a la profundidad del análisis de ciertos textos.

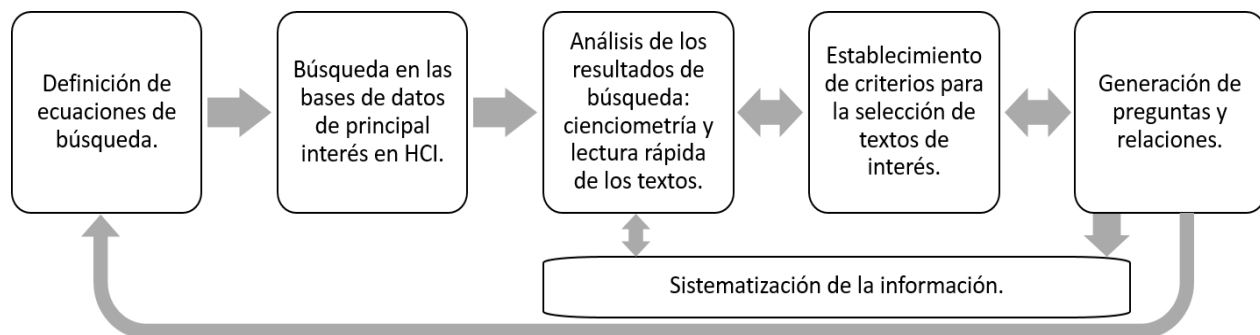


Fig. 1. Proceso metodológico para hacer la búsqueda

Fuente: elaboración propia.

### 3. RESULTADOS Y DISCUSIÓN

#### 3.1 Ecuaciones de búsqueda, bases de datos y criterios de selección

La primera fase consistió en la definición de los términos principales de búsqueda, con base en los grandes temas de investigación: i) gestos, ii) reconocimiento de gestos, iii) interacción.

Con esta información como insumo, se seleccionaron los sinónimos y otros términos relacionados. Después de definidas las ecuaciones de búsqueda, se eligieron las bases de datos para la revisión de artículos, para lo cual se utilizó como fuente las páginas web de Special Interest Group on Computer-Human Interaction (SIGCHI) [21], International Federation for Information Processing (IFIP) [22] y IEEE Systems, Man and Cybernetics Society IEEE (SMC) [23], por ser los principales referentes en el área de HCI.

La Tabla 1 resume las ecuaciones de búsqueda más significativas y los resultados obtenidos en las bases de datos de mayor publicación y consulta.

De acuerdo con la información de la Tabla 1, se infiere que la cantidad de resultados en *Springer* es significativamente mayor que en ACM o IEEE, debido a que tiene repositorios de áreas del conocimiento diversas: psicología,

medicina, estudios culturales y sociales, ciencias de la computación, entre otras.

Cuando la búsqueda se filtra por fechas en *Springer*, se encuentra que antes de los años noventa la mayoría de los resultados obedecen a las áreas de medicina y psicología; pero, a partir de dicha década, se concentran en ciencias de la computación. Sobre los resultados en IEEE y ACM, se observa que IEEE tiene más resultados en temas técnicos: reconocimiento de gestos y modelado de gestos; mientras que ACM arroja más resultados, comparativamente, cuando se incluyen palabras clave más asociadas a lo humano como “interacción” o “cultura”.

Adicionalmente, en los medios de divulgación se encuentran: ACM Interactions [24], Communications of the ACM [25], Interaction Design Foundation [26]; y, en Latinoamérica, la revista de diseño de interacción *Faz* [27], así como redes sociales y plataformas de video usadas por conferencias destacables como Conference on Human Factors in Computing Systems (CHI) [28], Tangible, Embedded and Embodied Interactions (TEI) [29] o, de la comunidad latinoamericana, Conferencia Latinoamericana en Interacción Humano Computador (CLIHIC) [30] y Conferencia Latinoamericana en Informática (CLEI).

Tabla 1. Ecuaciones de búsqueda principales, junto con las fechas y la cantidad de resultados encontrados en las principales bases de datos de consulta. Fuente: elaboración propia.

Ecuación de búsqueda	Variable	ACM	Springer	IEEE Xplore
Gesture AND Based AND Interaction	Rango de fechas de publicación	1988-2019	1929-2019	1989-2019
	Cantidad de resultados	2292	63055	2550
Gestural AND Interaction	Rango de fechas de publicación	1980-2019	1960-2019	1987-2019
	Cantidad de resultados	4762	7427	267
Gesture AND Recognition	Rango de fechas de publicación	1980-2019	1853-2019	1989-2019
	Cantidad de resultados	1366	59047	9518
Gesture AND Modelling	Rango de fechas de publicación	1970-2019	1850-2019	1978-2019
	Cantidad de resultados	1074	86616	4391
Gesture AND Culture	Rango de fechas de publicación	1991-2019	1850-2019	1997-2018
	Cantidad de resultados	118	62962	57

El ordenamiento para la aproximación a los textos fue inicialmente por relevancia, posteriormente, por número de citas y, finalmente, del más reciente al más antiguo. La selección de textos para una lectura más detallada dependió de la información encontrada en la primera aproximación, para lo cual se tuvieron en cuenta factores como: conferencia o revista en la que se publicó, relevancia de los autores (número de publicaciones y citas), rigurosidad en la descripción del texto en el resumen (por ejemplo, si presenta un método, describe una herramienta, algoritmo, o experimento con usuarios). Adicionalmente, se hizo una búsqueda específica en revistas en español (*Faz* [27]) y en fuentes de publicación latinoamericanas.

### 3.2 El gesto como medio de interacción

Los gestos son, en esencia, un medio de comunicación. Si bien se suelen considerar como naturales, de acuerdo con Norman [16], las expresiones gestuales no son naturales ni innatas ni fáciles de aprender, comprender o recordar. Más aún, los gestos son expresiones culturales [31] y muchos de ellos pueden ser efímeros para adaptarse a condiciones particulares del

contexto espacial o temporal en el que ocurren [16]. Así, existen dos escenarios posibles: i) desarrollar estrategias computacionales para el reconocimiento de gestos culturalmente aceptados-apropiados (adaptación) o ii) proponer gestos que sean reconocidos por los dispositivos computacionales y que puedan ser aprendidos y apropiados por las personas (estandarización). El primer escenario es el más complejo, por la diversidad y variabilidad de los gestos, entre una cultura y otra. El segundo escenario es el más común, y ha permitido la estandarización de gestos para la interacción humano-computador, por entidades formales o por convención a través de la popularización de su uso [32].

La Tabla 2 presenta una comparación de los dos enfoques, en cuanto a ventajas y desventajas. En ambos casos, la comprensión del gesto es muy amplia y aplica para diferentes partes del cuerpo: estos faciales [33], con las manos [5], [34], [35], con los dedos [4], [5], [20], o de cuerpo entero [8], [36]. En algunos casos, la complejidad técnica de reconocer la secuencia de acciones que definen un gesto ha llevado a que se utilicen objetos que restrinjan la interacción [11], [18].

Tabla 2. Ventajas y desventajas comparativas entre las dos estrategias utilizadas en HCI para el uso de gestos como medios de interacción. Fuente: adaptada de [32].

	Adaptación	Estandarización
<b>Ventajas</b>	<ul style="list-style-type: none"> <li>• Ergonomía. Comodidad y eficacia durante la interacción.</li> <li>• La interacción se puede personalizar, de manera que considera las particularidades de cada persona.</li> <li>• Flexibilidad, respecto a los cambios que puedan tener los gestos.</li> <li>• Reconocimiento contextual.</li> </ul>	<ul style="list-style-type: none"> <li>• De fácil reconocimiento, culturas de masas.</li> <li>• Permite transmitir el aprendizaje entre varias aplicaciones.</li> <li>• Reutilización.</li> <li>• Interoperabilidad.</li> <li>• Simplificación en las etapas de desarrollo y mantenimiento de aplicaciones.</li> </ul>
<b>Desventajas</b>	<ul style="list-style-type: none"> <li>• Aislamiento. Difícil de interoperar y transferir aprendizaje.</li> <li>• Evaluación compleja.</li> <li>• Dificultad en crear aplicaciones masivas.</li> </ul>	<ul style="list-style-type: none"> <li>• Desconocimiento de las particularidades que cada persona le da a un gesto.</li> <li>• Reconocimiento descontextualizado.</li> <li>• Gestos rígidos: no contemplan que los gestos pueden ser efímeros.</li> </ul>

### 3.3 Reconocimiento de gestos

El proceso de reconocimiento de los gestos requiere seguir los fundamentos de HCI: definir modelos de expresión, establecer un rango posible de acciones, restringir el espectro de interacciones, reconocer respuestas —deseadas e indeseadas— del sistema [16], [37].

Así mismo, exige afrontar nuevos desafíos que se encuentran en todas las etapas del proceso de reconocimiento de gestos, en este texto, comprendido en dos grandes etapas: entrenamiento y clasificación/regresión (Fig. 2).

El reconocimiento del gesto se aplica tanto para procesos de clasificación, en los que se etiqueta un conjunto de acciones del usuario, como para procesos de regresión, en los que se anticipa una acción consecuente, a partir de unas acciones previas.

En los dos casos, el proceso de entrenamiento ocurre en cuatro etapas, resumidas en la Tabla 3: captura, pre-procesamiento, selección y extracción de características y aprendizaje. La captura de gestos se suele hacer con diferentes tipos de sensores, entre los que se encuentran: ópticos (cámaras como la PlayStation Eye y cámaras de profundidad como Kinect o LeapMotion) [36], [38], [39]; acelerómetro (SteamVR tracker, Wii Mote); capacitivos (touchpad, pantallas táctiles de dispositivos móviles); giróscopos y electrodos para detección de movimientos musculares (EMG-Myo-, EEG-Emotiv-) [37].

Además de la selección de sensores, el diseño de la captura de gestos debe responder a cuestiones como: ¿cada cuánto tiempo se captura un dato?, ¿cuál es la resolución requerida para el convertidor analógico-digital?, ¿cuánto almacenamiento se requiere para guardar los datos crudos que provee el sensor?

El preprocesamiento, así como la extracción y selección de características, dependen de la aplicación particular y de la forma en que se reconozcan los gestos (por ejemplo, [40]-[43]). Estas etapas varían según el método de captura y la técnica de reconocimiento y pueden utilizar desde procesos estadísticos (análisis por componentes principales *PCA* [44] o discriminantes de Fisher), hasta algoritmos propios del contexto de aplicación como el de Haar [45], [46].

Algunas de las cuestiones que deben ser respondidas en estas etapas son: ¿qué ventana de tiempo es adecuada para preprocesar un conjunto de datos crudos?, ¿cómo se filtran los datos crudos?, ¿de qué manera se pueden comprimir los datos crudos para mantener la información más relevante, sin demandar demasiado espacio de memoria?, ¿se preprocesa el gesto luego de capturarlo completamente (*offline*) o se va preprocesando en tiempo real, a medida que la persona lo hace (*online*)?, ¿cómo se representa computacionalmente el gesto? Y, si se pierde un segmento de la información, ¿cómo se conserva el gesto?

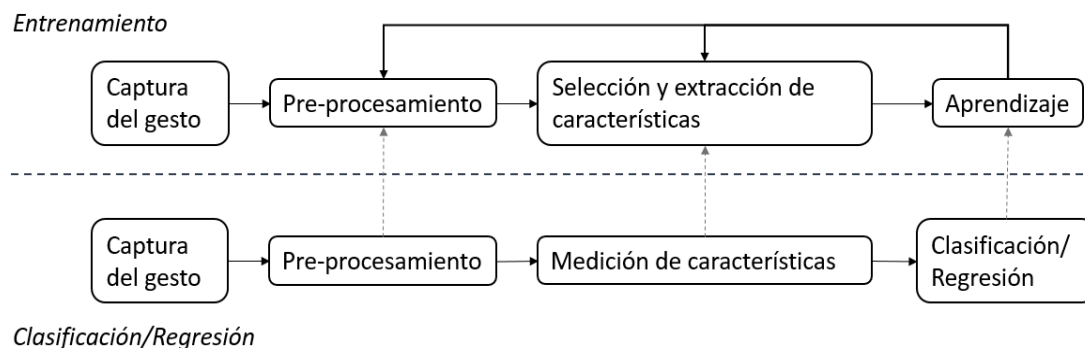


Fig. 2. Etapas del reconocimiento de gestos. Fuente: elaboración propia.

Tabla 3. Resumen de las diferentes etapas a considerar en el proceso de reconocimiento de gestos, incluidas las cuestiones que se deben resolver en cada etapa, técnicas y herramientas. Fuente: elaboración propia.

<b>Etapas</b>	<b>Cuestiones por resolver</b>	<b>Técnicas y herramientas</b>
Captura	<ul style="list-style-type: none"> <li>• Método de captura: sensores</li> <li>• Frecuencia de muestreo</li> <li>• Resolución de la captura</li> <li>• Memoria requerida</li> <li>• Formato de los datos crudos</li> </ul>	<b>Sensores:</b> <ul style="list-style-type: none"> <li>• Ópticos 2D o de profundidad</li> <li>• Acelerómetros</li> <li>• Capacitivos</li> <li>• Electrodo</li> </ul>
Preprocesamiento	<ul style="list-style-type: none"> <li>• Conversión de los datos crudos</li> <li>• Preprocesamiento <i>offline</i> u <i>online</i></li> <li>• Representación computacional del gesto</li> <li>• Tolerancia a errores en la captura</li> </ul>	<ul style="list-style-type: none"> <li>• Conocimiento del gesto: experticia que informa el proceso de reconocimiento</li> </ul>
Selección y extracción de características	<ul style="list-style-type: none"> <li>• ¿Se requiere hacer selección o extracción de características?</li> <li>• Escoger el algoritmo de selección o extracción de características de acuerdo con el enfoque.</li> </ul>	<b>Extracción:</b> <ul style="list-style-type: none"> <li>• PCA</li> <li>• Haar</li> <li>• Discriminantes lineales de Fischer</li> <li>• ...</li> </ul> <b>Selección:</b> <ul style="list-style-type: none"> <li>• Filtro</li> <li>• <i>Wrapper</i></li> <li>• Embebido</li> </ul>
Aprendizaje	<ul style="list-style-type: none"> <li>• Selección del algoritmo de aprendizaje</li> <li>• Selección de la técnica de evaluación/validación</li> </ul>	<b>Aprendizaje:</b> <ul style="list-style-type: none"> <li>• Modelos Ocultos de Markov (HMM)</li> <li>• Redes neuronales</li> <li>• Freeman Chain Code</li> <li>• Máquinas de Soporte Vectorial</li> <li>• Aprendizaje profundo</li> <li>• ...</li> </ul> <b>Evaluación:</b> <ul style="list-style-type: none"> <li>• Validación cruzada</li> <li>• Matriz de confusión</li> <li>• Sensibilidad, especificidad, exactitud</li> </ul>

Para el aprendizaje de máquina y la etapa de clasificación o regresión, se han utilizado diferentes técnicas, entre las que destacan: Modelos Ocultos de Markov (*HMM*) [15], [33], [34], [47], Redes Neuronales [34], Freeman Chain Code [44], Máquinas de Vectores de Soporte [48]-[50], evolución neuronal [35] y aprendizaje profundo [51], [52].

Para revisar las tendencias sobre algoritmos de reconocimiento utilizados en la clasificación o regresión de gestos, se hizo una búsqueda filtrada por la fecha de publicación, en la cual se escogieron solo

artículos del presente año: 2019. Asimismo, se utilizó la ecuación de búsqueda: “Gesture AND (Classification OR Regression OR Recognition)”. Luego, se concatenaron los textos de título, resumen y palabras clave de los resultados y se contaron las palabras para estimar cuáles son los algoritmos más utilizados actualmente.

Los hallazgos muestran que el aprendizaje profundo (*deep learning*) y, en particular, las redes neuronales convolucionales son los algoritmos utilizados con mayor frecuencia (ver Tabla 4).



Tabla 4. Palabras más frecuentes en relación con los algoritmos de aprendizaje, en los títulos, resúmenes y palabras clave resultantes de la ecuación “*Gesture AND (Classification OR Regression OR Recognition)*”

Fuente: elaboración propia.

Palabra	Conteo	Palabra	Conteo
Recognition	1305	Clustering	53
Learning	1078	LSTM	44
		Tree	38
<b>Deep</b>	<b>517</b>	Logistic	38
<b>Network</b>	<b>494</b>	Unsupervised	37
<b>Neural</b>	<b>476</b>	Bayesian	32
...	...	SVM	28
<b>Convolutional</b>	<b>217</b>	Fuzzy	28
CNN	100	HMM	24
Supervised	71	Heuristic	10

Adicionalmente, se encontró también la existencia de palabras frecuentes como *clustering* y *unsupervised*, asociadas a técnicas no supervisadas; o *tree* y *fuzzy*, asociadas a árboles de decisión y lógica difusa como técnicas no estadísticas.

El análisis sobre las tendencias finalizó con la selección de la fuente de búsqueda *Springer*, debido a que tenía una mayor cantidad de resultados en las publicaciones referentes a reconocimiento y a modelado.

Como indicador se obtuvo la cantidad de publicaciones en las que se mencionara el algoritmo o técnica de clasificación, de acuerdo con las siguientes ecuaciones de búsqueda: i) Deep Learning: *Gesture AND (“deep learning” OR “convolutional neural network”)*, ii) SVM: *Gesture AND (“Support Vector Machine” OR “SVM” OR “SVR”)*, iii) Redes neuronales: *Gesture AND “neural networks” AND NOT “Deep” AND NOT “convolutional”*, iv) HMM: *Gesture AND (“hidden markov model” OR “HMM”)*.

Las publicaciones incluidas en el conteo fueron separadas por año y limitadas a las áreas de ingeniería y ciencias de la computación. Los resultados se muestran en la Fig. 2.

### 3.4 Aplicaciones y desafíos

En las áreas de aplicación de las interacciones basadas en gestos que se encuentran más a menudo [5], sobresalen la de robótica [15], [33], [47], [53], en particular el campo de la robótica suave; las tareas recurrentes en HCI como seleccionar, maximizar, arrastrar, copiar, pegar [54]; la preservación cultural [44]; la computación incluyente que involucra grupos poblacionales particulares como niños, adultos mayores, personas con capacidades limitadas o movilidad reducida; o los videojuegos [55]-[57].

En esos campos, se detectaron los siguientes desafíos de interés actual: i) *Segmentación* [58]. En aplicaciones en las que un gesto siga al siguiente, ¿cómo reconocer cuándo acaba un gesto e inicia el otro?, ¿cómo separar un gesto del contexto en el que ocurre? (del fondo). ii) *Concurrencia*. Si un gesto es capturado con múltiples sensores, del mismo o diferente tipo, o usando diferentes máquinas, ¿cómo fusionar los datos?; si el sistema de cómputo es utilizado por múltiples personas a la vez, ¿cómo diferenciar los gestos de una persona respecto a otra?, ¿existen los gestos colectivos, que son ejecutados por más de una persona?

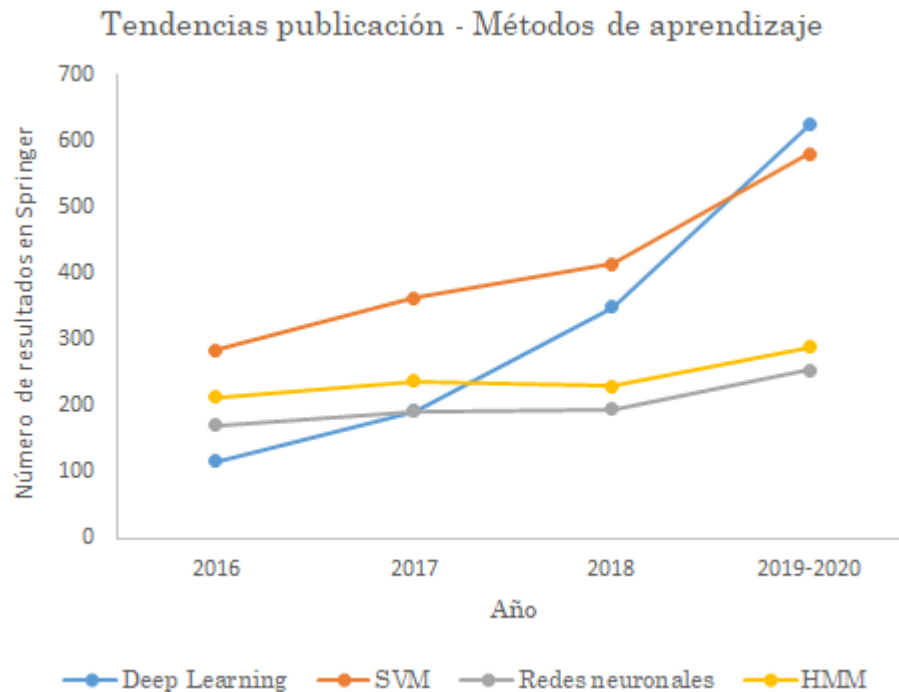


Fig. 3. Tendencias en el número de publicaciones asociadas a gestos, según el año y el tipo de método de aprendizaje automático. Fuente: elaboración propia.

En caso tal, ¿cómo reconocer dichos gestos? iii) *Experiencia de usuario*. ¿Cómo garantizar que el usuario puede completar una tarea con un gesto?, ¿cómo realimentar al usuario cuando no puede completar una acción a través de un gesto?, ¿qué herramientas se proveen al usuario en caso de que no pueda hacer un gesto?, ¿cómo enseñarle al usuario a hacer gestos que sean reconocidos por la máquina? iv) *Estudios culturales*. ¿Cómo representar computacionalmente los gestos?, ¿es posible estandarizar esa representación?, ¿cómo afecta el contexto cultural el proceso de reconocimiento computacional de gestos?, ¿de qué manera desarrollar estudios culturales relacionados con las expresiones gestuales de una comunidad, desde el área de HCI? v) *Adaptación*. ¿Cómo diseñar máquinas que no solo reconozcan gestos específicos, sino que puedan adaptarse a los gestos de sus usuarios?, ¿de qué forma se pueden adaptar los algoritmos de reconocimiento de gestos para que sea posible que se modifiquen cuando los gestos de las personas cambian?

#### 4. CONCLUSIONES

El tema de interacciones basadas en gestos se percibe como reciente. Sin embargo, la interacción de los humanos con herramientas, máquinas e instrumentos musicales, a través de gestos, es innata.

No por esto se puede considerar que las interfaces basadas en gestos son naturales; si bien la comunicación no verbal es parte de la naturaleza humana, los gestos con los que se lleva a cabo varían de una cultura a otra, son efímeros, pueden cambiar con el contexto o el tiempo.

De otro lado, la investigación en gestos como medio de interacción con los computadores se ha centrado en la dirección humano a máquina; esto es, cómo los computadores reconocen gestos para que el sistema responda a ciertas intenciones o acciones del usuario.

Sin embargo, queda abierta la posibilidad de preguntarse acerca de la comunicación en sentido contrario: ¿pueden las máquinas expresar gestos por los que los humanos sientan empatía (o

apatía)? En temas de inteligencia artificial, se hace referencia a la solución de problemas por sistemas artificiales, lo que ha derivado en sistemas que resuelven problemas como el juego de GO. Sin embargo, en problemas en los que se requiere un componente corporal y de empatía, como en un juego de póquer, ¿cuál es el rol de la máquina?, ¿será posible resolver este tipo de problemas y que la máquina “le gane” al humano?

Así mismo, la interacción humano-computador también ha estudiado los gestos como un medio para evaluar la interacción entre el humano y la máquina, de acuerdo con las expresiones faciales y corporales del usuario durante la interacción.

Finalmente, la investigación y desarrollo tecnológico de métodos y herramientas que permitan la comunicación entre el humano y el computador, a través de gestos, es un tema vigente, de interés tanto para la comunidad académica como para la industria. El trabajo es aún incipiente, pero ha facultado la formulación de diversas preguntas de investigación, algunas mencionadas en el presente artículo. Es de gran importancia una visión holística de las interacciones basadas en gestos, que permita afrontar desafíos técnicos, también desde dimensiones culturales, sociales y psicológicas.

## 5. AGRADECIMIENTOS

Agradecemos a la Universidad Militar Nueva Granada y a la Institución Universitaria Politécnico Grancolombiano, por permitirnos el tiempo para el desarrollo de la revisión y escritura de este artículo.

## 6. REFERENCIAS

- [1] W. Grimshaw, “An etymological dictionary or analysis of the English language”, Trieste Publishing, 2018.
- [2] Real Academia Española, “«Diccionario de la lengua española» - Edición del Tricentenario», «Diccionario de la lengua española» - Edición del Tricentenario.
- [3] E. de Lera y M. Garreta-Domingo, “10 heurísticos emocionales - pautas para evaluar la dimensión afectiva de los usuarios de forma fácil y económica,” *Rev. Faz*, no. 2, pp. 68–81, Jul. 2008. Disponible en: [URL](#)
- [4] H. Lu and Y. Li, “Gesture On: Enabling Always-On Touch Gestures for Fast Mobile Access from the Device Standby Mode,” en *Proceedings of the 133rd Annual ACM Conference on Human Factors in Computing Systems - CHI '15*, Seoul, 2015, pp. 3355–3364. <https://doi.org/10.1145/2702123.2702610>
- [5] L. Chen, F. Wang, H. Deng, y K. Ji, “A Survey on Hand Gesture Recognition,” en *2013 International Conference on Computer Sciences and Applications*, Wuhan, 2013, pp. 313–316. <https://doi.org/10.1109/CSA.2013.79>
- [6] E. McAweeney, H. Zhang, y M. Nebeling, “User-Driven Design Principles for Gesture Representations,” en *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems - CHI '18*, Montreal, 2018, pp. 1-13. <https://doi.org/10.1145/3173574.3174121>
- [7] R. Vatavu, L. Anthony, y J. O. Wobbrock, “Gestures as point clouds: A precognizer for user interface prototypes,” en *Proceedings of the 14th ACM international conference on Multimodal interaction - ICMI '12*, California, 2012, pp. 273 – 280. <https://doi.org/10.1145/2388676.2388732>
- [8] S. Piana, A. Staglianò, F. Odone, y A. Camurri, “Adaptive Body Gesture Representation for Automatic Emotion Recognition,” *ACM Trans. Interact. Intell. Syst.*, vol. 6, no. 1, pp. 1–31, Mar. 2016. <https://doi.org/10.1145/2818740>
- [9] M. Denisa, A. Gams, A. Ude, y T. Petric, “Learning Compliant Movement Primitives Through Demonstration and Statistical Generalization,” *IEEE/ASME Trans. Mechatronics*, vol. 21, no. 5, pp. 2581–2594, Oct. 2016. <https://doi.org/10.1109/TMECH.2015.2510165>

- [10] M. Neff, M. Kipp, I. Albrecht, y H.-P. Seidel, "Gesture modeling and animation based on a probabilistic re-creation of speaker style," *ACM Trans. Graph.*, vol. 27, no. 1, pp. 1–24, Mar. 2008.  
<https://doi.org/10.1145/1330511.1330516>
  - [11] I. Poupyrev, T. Nashida, y M. Okabe, "Actuation and tangible user interfaces: the Vaucanson duck, robots, and shape displays," in *Proceedings of the 1st international conference on Tangible and embedded interaction - TEI '07*, Baton Rouge, Louisiana, 2007, pp. 205–212.  
<https://doi.org/10.1145/1226969.1227012>
  - [12] R. Arnheim y D. McNeill, "Hand and Mind: What Gestures Reveal about Thought by David McNeill," *Leonardo*, vol. 27, no. 4, pp. 358, 1994.  
<https://doi.org/10.2307/1576015>
  - [13] C. S. Montero, J. Alexander, M. T. Marshall, y S. Subramanian, "Would you do that?: understanding social acceptance of gestural interfaces," en *Proceedings of the 12th international conference on Human computer interaction with mobile devices and services - MobileHCI '10*, Lisbon, 2010, pp. 275–278.  
<https://doi.org/10.1145/1851600.1851647>
  - [14] M. Rehm, N. Bee, y E. André, "Wave like an Egyptian: accelerometer based gesture recognition for culture specific interactions," in *BCS-HCI '08 Proceedings of the 22nd British HCI Group Annual Conference on People and Computers: Culture, Creativity, Interaction*, Liverpool, 2008, pp. 13–22. Disponible en: <URL>
  - [15] P. Trigueiros, F. Ribeiro, y L. P. Reis, "Generic system for human-computer gesture interaction," en *2014 IEEE International Conference on Autonomous Robot Systems and Competitions (ICARSC)*, Espinho, 2014, pp. 175–180.  
<https://doi.org/10.1109/ICARSC.2014.6849782>
  - [16] D. A. Norman, "Natural user interfaces are not natural," *Interactions*, vol. 17, no. 3, pp. 6–10, May 2010.  
<https://doi.org/10.1145/1744161.1744163>
  - [17] H. Ishii, "Tangible bits: beyond pixels," en *Proceedings of the 2nd international conference on Tangible and embedded interaction - TEI '08*, Bonn, 2008, pp. 15–25.  
<https://doi.org/10.1145/1347390.1347392>
  - [18] H. Ishii, "The tangible user interface and its evolution," *Commun. ACM*, vol. 51, no. 6, pp. 32–36, Jun. 2008.  
<https://doi.org/10.1145/1349026.1349034>
  - [19] D. Holman y R. Vertegaal, "Organic User Interfaces: Designing Computers in Any Way, Shape, or Form," *Commun. ACM*, vol. 51, no. 6, pp. 48–55, Jun, 2008.  
<https://doi.org/10.1145/1349026.1349037>
  - [20] N. Breslauer, I. Galić, M. Kukec, y I. Samardžić, "Leap Motion Sensor for Natural User Interface," *Teh. Vjesn. - Tech. Gaz.*, vol. 26, no. 2, pp. 560–565, Apr. 2019.  
<https://doi.org/10.17559/TV-20181012093055>
  - [21] Association for Computing Machinery SIGCHI, 2019. Disponible en: <URL>
  - [22] Interaction Design Foundation, "International Federation for Information Processing", 2019. Disponible en: <URL>
  - [23] Institute of Electrical and Electronics Engineers, Man, and Cybersecurity Society, 2019 en *IEEE international conference on systems, man, and cybernetics SMC2019*, Bari, 2019. Disponible en: <URL>
  - [24] Association for Computing Machinery, «Interactions», 2019. Disponible en: <URL>
  - [25] Association for Computing Machinery, «Communications of the ACM», 2019. Disponible en: <URL>
  - [26] Interaction Design Foundation, *Interaction Design Foundation*. 2019. Disponible en: <URL>
  - [27] Revista FAZ, «FAZ - Revista de diseño de interacción», no. 9. Aug. 2016. Disponible en: <URL>
  - [28] Association for Computing Machinery ACM SIGCHI, "CHI 2019" en *Conference of Human-Computer Interaction. CHI*, Glasgow, 2019. Disponible en: <URL>
  - [29] Association for Computing Machinery SIGCHI "TEI 2019" en *13 international conference on tangible, Embedded, and embodied interactions*, Tempe, Arizona, 2019. Disponible en: <URL>
  - [30] Latin American, "CLIHIC 2019", en *IX Latin American Conference on Human Computer Interaction, LAIHC*, Panamá, 2019. Disponible: <URL>
  - [31] D. Mauney, J. Howart, A. Wirtanen, y M. Capra, «Diferencias y similitudes culturales en gestos definidos por el usuario para interfaces en pantallas táctiles», *Revista FAZ - Gestuales, tangibles y de cuerpo entero: Nuevas interacciones*, vol. 4, pp. 16–25, Oct. 2010. Disponible en: <URL>
  - [32] M. Bobillier-Chaumon, S. Carvallo, F. Tarpin Bernard y J. Vancherand Revel "Adapter ou
- [130] TecnoLógicas, ISSN-p 0123-7799 / ISSN-e 2256-5337, Vol. 22, edición especial, noviembre de 2019, pp. 119–132

- uniformiser les interactions personnes-systèmes? to adapt or standardize the human-computer interactions?”, *Rev. D'Interaction Homme-Mach.*, vol. 6, no 2, pp. 91-129, 2005. Disponible en: [URL](#)
- [33] J. G. Bueno, M. González-Fierro, C. Balaguer, y L. Moreno, “Facial gesture recognition using active appearance models based on neural evolution,” en *2012 7th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, pp. 133–134. Massachusetts, 2012, Disponible en: [URL](#)
- [34] S. Bhowmick, A. K. Talukdar, y K. K. Sarma, “Continuous hand gesture recognition for English alphabets,” en *2015 2nd International Conference on Signal Processing and Integrated Networks (SPIN)*, Noida, 2015. pp. 443–446.  
<https://doi.org/10.1109/SPIN.2015.7095264>
- [35] Y. Li *et al.*, “Hand Gesture Recognition and Real-time Game Control Based on A Wearable Band with 6-axis Sensors,” en *2018 International Joint Conference on Neural Networks (IJCNN)*, Rio de Janeiro, 2018, pp. 1–6.  
<https://doi.org/10.1109/IJCNN.2018.8489743>
- [36] B. W Hwang, S. Kim, and S. W. Lee, “A Full-Body Gesture Database for Automatic Gesture Recognition,” in *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, Southampton, 2006, pp. 243–248.  
<https://doi.org/10.1109/FGR.2006.8>
- [37] L. R. Ruiz, F. De la Rosa R., y J. T. Hernandez, “Platform integrating interactive applications with gesture-based interaction,” en *2012 XXXVIII Conferencia Latinoamericana En Informatica (CLEI)*, Medellín, 2012, pp. 1–9.  
<https://doi.org/10.1109/CLEI.2012.6427129>
- [38] P. Ponsa, C. Urbina, C. Manresa-Yee, y R. Vilanova, «Estudio de Usabilidad de una Interfaz Gestual Basada en Visión», *Revista FAZ*, vol. 8, pp. 99-119, 2015. Disponible en: [URL](#)
- [39] O. Patsadu, C. Nukoolkit, and B. Watanapa, “Human gesture recognition using Kinect camera,” in *2012 Ninth International Conference on Computer Science and Software Engineering (JCSSE)*, Bangkok, 2012. pp. 28–32.  
<https://doi.org/10.1109/JCSSE.2012.6261920>
- [40] S. E. Nope, H. Loaiza, and E. Caicedo, “Modelo Bio-inspirado para el Reconocimiento de Gestos Usando Primitivas de Movimiento en Visión,” *Rev. Iberoam. Automática e Informática Ind. RIAI*, vol. 5, no. 4, pp. 69–76, Oct. 2008.  
[https://doi.org/10.1016/S1697-7912\(08\)70179-1](https://doi.org/10.1016/S1697-7912(08)70179-1)
- [41] M. B. Kaaniche y F. Bremond, “Gesture recognition by learning local motion signatures,” en *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, 2010, pp. 2745–2752.  
<https://doi.org/10.1109/CVPR.2010.5539999>
- [42] L.-P. Morency, A. Quattoni, and T. Darrell, “Latent-Dynamic Discriminative Models for Continuous Gesture Recognition,” en *2007 IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, 2007, pp. 1–8.  
<https://doi.org/10.1109/CVPR.2007.383299>
- [43] H.-I. Suk, B.-K. Sin, and S.-W. Lee, “Hand gesture recognition based on dynamic Bayesian network framework,” *Pattern Recognit.*, vol. 43, no. 9, pp. 3059–3072, Sep. 2010.  
<https://doi.org/10.1016/j.patcog.2010.03.016>
- [44] J. Xu, S. Sun, Y. Shi, y Z. Dong, “Implementation of Digital Chime-bell Interaction System Driven by Hand Gesture,” en *2008 Third International Conference on Pervasive Computing and Applications*, Alexandria, 2008, pp. 264–268.  
<https://doi.org/10.1109/ICPCA.2008.4783590>
- [45] A. Królak, “Use of Haar-like features in vision-based human-computer interaction systems,” en *2012 Joint Conference New Trends In Audio & Video And Signal Processing: Algorithms, Architectures, Arrangements And Applications (NTAV/SPA)*, Lodz, 2012, pp. 139–142. Disponible en: [URL](#)
- [46] S. S. Rautaray y A. Agrawal, “Interaction with virtual game through hand gesture recognition,” en *2011 International Conference on Multimedia, Signal Processing and Communication Technologies*, Aligarh, 2011, pp. 244–247.  
<https://doi.org/10.1109/MSPCT.2011.6150485>
- [47] S.-O. Shin, D. Kim, y Y.-H. Seo, “Controlling Mobile Robot Using IMU and EMG Sensor-Based Gesture Recognition,” en *2014 Ninth International Conference on Broadband and Wireless Computing, Communication and Applications*, Guangdong, 2014, pp. 554-557.  
<https://doi.org/10.1109/BWCCA.2014.145>
- [48] M. P. Tarvekar, “Hand Gesture Recognition System for Touch-Less Car Interface Using



- Multiclass Support Vector Machine,” en *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, Madurai, 2018, pp. 1929–1932.  
<https://doi.org/10.1109/ICCONS.2018.8663003>
- [49] K. Feng y F. Yuan, “Static hand gesture recognition based on HOG characters and support vector machines,” en *2013 2nd International Symposium on Instrumentation and Measurement, Sensor Network and Automation (IMSNA)*, pp. 936–938, Toronto, 2013.  
<https://doi.org/10.1109/IMSNA.2013.6743432>
- [50] J. Oh, T. Kim, y H. Hong, “Using Binary Decision Tree and Multiclass SVM for Human Gesture Recognition,” en *2013 International Conference on Information Science and Applications (ICISA)*, Suwon, 2013, pp. 1–4.  
<https://doi.org/10.1109/ICISA.2013.6579388>
- [51] W. Nan, Z. Zhigang, L. Huan, M. Jingqi, Z. Jiajun, y D. Guangxue, “Gesture Recognition Based on Deep Learning in Complex Scenes,” en *2019 Chinese Control And Decision Conference (CCDC)*, Nanchang, 2019, pp. 630–634.  
<https://doi.org/10.1109/CCDC.2019.8833349>
- [52] U. Cote-Allard *et al.*, “Deep Learning for Electromyographic Hand Gesture Signal Classification Using Transfer Learning,” *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 27, no. 4, pp. 760–771, Apr. 2019.  
<https://doi.org/10.1109/TNSRE.2019.2896269>
- [53] T. Obo, R. Kawabata, y N. Kubota, “Cooperative Human-Robot Interaction Based on Pointing Gesture in Informationally Structured Space,” en *2018 World Automation Congress (WAC)*, Stevenson, 2018, pp. 1–5.  
<https://doi.org/10.23919/WAC.2018.8430388>
- [54] C. A. Díaz-León, E. M. Hincapié-Montoya, E. A. Guirales-Arredondo, and G. A. Moreno-López, “Diseño y desarrollo de un sistema de interacción para su implementación en un aula de clase inteligente,” *Rev. EIA*, vol. 13, no. 26, pp. 95–109, Jul. 2016.  
<https://doi.org/10.24050/reia.v13i26.666>
- [55] L. Vázquez, A. Martínez, y G. López, “Videojuego serio como apoyo a la estimulación temprana del pensamiento matemático”, *Revista FAZ*, vol. 9, pp. 13–31, Jul. 2016. Disponible en: [URL](#)
- [56] C. A. Castillo-Benavides, L. F. García-Arias, N. D. Duque-Méndez, y D. A. Ovalle-Carranza, “IMU-Mouse: diseño e implementación de un dispositivo apuntador dirigido al desarrollo de interfaces adaptativas para personas con discapacidad física,” *TecnoLógicas*, vol. 21, no. 41, pp. 63–79, Jan. 2018.  
<https://doi.org/10.22430/22565337.727>
- [57] D. J. Botina-Monsalve, M. A. Domínguez-Vásquez, C. A. Madrigal-González, y A. E. Castro-Ospina, “Clasificación automática de las vocales en el lenguaje de señas colombiano,” *TecnoLógicas*, vol. 21, no. 41, pp. 103–114, Jan. 2018.  
<https://doi.org/10.22430/22565337.730>
- [58] J. Alon, V. Athitsos, Quan Yuan, y S. Sclaroff, “A Unified Framework for Gesture Recognition and Spatiotemporal Gesture Segmentation,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 9, pp. 1685–1699, Sep. 2009.  
<https://doi.org/10.1109/TPAMI.2008.203>